Deliberate System-Side Errors as a Potential Pedagogic Strategy for Exploratory Virtual Learning Environments

Citation for published version:
Alcorn, AM, Good, J & Pain, H 2013, Deliberate System-Side Errors as a Potential Pedagogic Strategy for Exploratory Virtual Learning Environments. in *Artificial Intelligence in Education - 16th International Conference, AIED 2013, Memphis, TN, USA, July 9-13, 2013. Proceedings.* pp. 483-492. https://doi.org/10.1007/978-3-642-39112-5_49

Digital Object Identifier (DOI):
10.1007/978-3-642-39112-5_49

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Artificial Intelligence in Education - 16th International Conference, AIED 2013, Memphis, TN, USA, July 9-13, 2013. Proceedings

Deliberate System-Side Errors as a Potential Pedagogic Strategy for Exploratory Virtual Learning Environments

Alyssa M. Alcorn[1], Judith Good[2], and Helen Pain[1]

[1] University of Edinburgh, School of Informatics, UK
{a.alcorn,helen.pain}@ed.ac.uk
[2] University of Sussex, School of Informatics, UK
j.good@sussex.ac.uk

Abstract. This paper describes an exploratory study of system-side errors (i.e. expectation- or rule-violations) in a *virtual environment* (VE), and the subsequent reactions of young children with *autism spectrum conditions* (ASC). Analysis of existing video from 8 participants interacting with the ECHOES VE showed that they frequently detected and reacted to system-side errors, engaging in social and communicative behaviours targeted by ECHOES. Detecting errors requires children to compare the VE's state to their "mental model" of its behaviour, determining where the two are discrepant. This is equivalent to learners identifying mistakes in their *own* knowledge and then re-aligning with the system-as-expert. This paper explores the implications of these results, proposing a taxonomy of discrepant event types, and discussing their location with respect to the learner and/or system. In addition to considering these results' significance for this user group and context, it relates the research to existing work that uses erroneous examples.

Keywords: Virtual environments, discrepancy, system error, learner error, learning, model, Autism, children, social communication, initiation, evaluation, HCI, design.

1 Introduction

Virtual learning environments and other adaptive systems have tended to focus their efforts on identifying and correcting errors in the learner's understanding or procedural knowledge. The knowledge shared by learner and system increases, with false or incomplete ("buggy") learner knowledge decreasing as a proportion of the total. Depending on the domain and the system, this may take the form of explicitly correcting steps in a learner's work or more subtly promoting relevant information and strategies (see [1] for a range of examples).

While the system's[1] domain knowledge must exceed the learner's in order to scaffold his/her progress, there is an important difference between an infallible system and one

[1] *System* is used in this paper as a generic term that encompasses adaptive learning environments, VEs, intelligent tutoring systems, serious games, and related projects that utilize technology for some teaching or practice function.

K. Yacef et al. (Eds.): AIED 2013, LNAI 7926, pp. 483–492, 2013.

that *overall* knows more, but makes occasional errors (as would a human teacher). If there are mistakes[2], they are generally presented as a deliberate teaching device, such as inviting learners to identify incorrect steps in worked mathematics examples [2, 3]. Occasional *system-side errors* that are *not* explicitly announced as problem-solving tasks may provide an opportunity for the learner to engage in metacognition, articulating his or her knowledge in order to address them. When errors constitute a relatively small proportion of the system-learner interactions, learners can take advantage of these metacognitive opportunities, and continue to benefit from the system's overall expertise.

This paper describes a study of system-side errors[3] in existing video data from the ECHOES virtual environment (VE). ECHOES was designed to help support young children with *autism-spectrum conditions* (ASC) to practice foundational social communication skills through exploratory play (see Section 2, and [4, 5])[4]. Its content, and thus the errors, are highly visual and focus on cause-and-effect relationships and patterns rather than factual knowledge. This cause-and-effect knowledge is never explicitly taught, but acquired over the course of the child's exploration. The system-side errors were a completely unintentional byproduct of the AI planner, rather than a deliberate design choice. Indeed, the characteristics of autism mean that expectation-violating aspects would generally be considered a poor, potentially upsetting choice for this user group (see Section 2). Nevertheless, the errors were highly effective in motivating children to engage the positive social communication behaviours that ECHOES tried to promote. In particular, children initiated to the human researcher and the ECHOES virtual character (VC) about the content of the system errors, sometimes explicitly indicating what *should* have happened instead (i.e. they were able to correct the system's error).

This error-detection process is inherently metacognitive [6], in that children had to compare their knowledge, expectations of, or predictions about the VE's contents and "rules" (i.e. their *mental model* of the system) to its actual behaviour, identifying mismatched aspects. This process of comparing models to identify *discrepancies* is arguably equivalent to learners identifying and correcting "bugs" in their own knowledge[5] by comparing themselves to an expert.

The ECHOES video analysis reported in this paper forms the basis for a more general discussion of discrepancy detection, including a taxonomy of discrepancy types and their possible sources in either a learner's mental model, or in a system. This paper explores the implications of these results for this particular user group and context, but also their relationship to existing work that uses erroneous examples.

2 The ECHOES Technology-Enhanced Learning Project

The ECHOES project developed a technology-enhanced learning environment targeted primarily at young children with ASC (aged 5-7 years), but with the potential to

---

[2] The terms 'error' and 'mistake' are used interchangeably in this paper.

[3] *Errors* do not mean error messages, or system freezes/crashes. They are errors in that the system violated its patterns of object or VC behaviour, or acted counter to activity goals.

[4] See `www.echoes2.org`

[5] More accurately, the learner corrects the mental model "for next time", as in most cases the process or interaction cannot actually be altered to reflect the correct action or information.

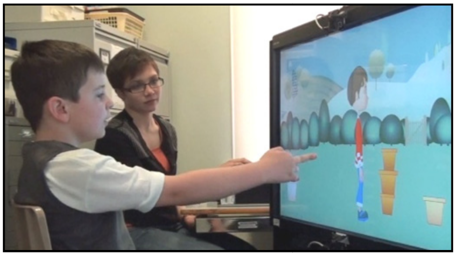
also be used by typically developing (TD) children [4, 5]. ECHOES includes a programme of game-like activities set in a "Magic Garden" VE, and was designed to support exploration and scaffolding of foundational social and communicative skills, such as turn-taking, and gaze- and point-following.

The ASC comprise a set of lifelong neuro-developmental conditions, characterised by notable and pervasive difficulties in communication and social interaction, plus the presence of repetitive behaviours and/or interests, sometimes manifested as a strong desire for routine and sameness [7]. Multiple VEs have already been developed to support children with ASC in learning specific skills (e.g. as discussed in [8]). The predictability, repeatability, and relative simplicity of VEs (compared to human-human interaction) are given as reasons why they are particularly suited to, and motivating for, this population, and may also be a useful research tool.

A young child using the ECHOES VE stands or sits in front of a 42" multi-touch screen, immersed in the visuals and sounds of the Magic Garden and physically involved in the interactions. ECHOES learning activities were developed with input from stakeholders, and draw strongly on educational and psychological theory. Activities encourage experimentation and play by deliberately introducing novel elements and behavioral fantasy, such as "pulling" on flower heads to transform them into bubbles or bouncy balls. The child has an autonomous, childlike VC (Andy) as a guide and playmate, demonstrating actions and offering encouragement. The underlying AI plans Andy's behaviour both deliberatively and in reaction to child actions (see [9]). Sound output is present, but dialogue is pre-recorded with no text-to-speech capability. There is also no capacity for speech recognition or sound input.

The system was designed such that children use it alongside an adult (researcher or teacher) who manages inter-activity transitions and gives limited system commands (such as for Andy to repeat an instruction) through a smaller, secondary screen (see Figure 1, Left). The adult does not direct the child's use of ECHOES. Instead, he/she plays an essential role in providing additional support for the child's complex communicative and emotional regulation needs (e.g. reformulating the VC's directions to include key instructional phrases or sign language familiar to that child). These cannot yet be met by an adaptive system in a rapid, robust, and appropriate fashion. Furthermore, an early ECHOES study [5] discovered that children frequently extended their interaction with the system to include the nearby adult, sharing their discoveries or seeking additional information.

28 children with ASC from four UK school sites participated in the summative evaluation of ECHOES (results in preparation). The goal was to assess a range of social and communication skills before, during, and after six to eight weeks of using the ECHOES environment. Children completed several 10-20 minute sessions with ECHOES per week, gradually encountering more complex material. Video data was the primary record of the child's communication and social behaviour. Each session was recorded by digital camcorder, positioned to capture the study environment



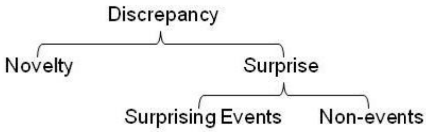


Fig. 1. (*Left)* ECHOES evaluation set-up. The researcher sits near the child, at the control monitor (not visible). (*Right)* Taxonomy of discrepancy types and their relationships, based on [10].

(see Figure 1, Left). This paper re-analyses a subset of the evaluation video data in order to explore the type and quantity of discrepancies and child reactions that were present (see [10]).

3 Types of Discrepancy

The ECHOES system included a range of errors in which the VE or VC appeared to make mistakes, compared to the pattern established by previous interactions. Participating children perceived many of these occurrences as *discrepant* from their expectations. In the current analysis, *discrepancy* has a child-centred definition. It is not an inherent property of mental models or environments, but exists via the process of the child making a comparison and detecting a mismatch between the mental model of the environment and some current aspect; if the child makes no comparison or fails to detect a difference, *then no discrepancy is present,* because the child's mental model agrees with the environment's current behaviour.

Many mistakes are examples of *surprise*— where something is known about the current aspect (it is part of the mental model), but it does not behave as expected. *Surprising events* include aspects that are present, but whose appearance or behavior is not as expected or predicted (i.e. in accordance with the mental model). In a *non-event,* some aspect of the environment violates expectations by unexpectedly or unpredictably being absent, being inactive or unresponsive, or failing to occur. A final type of discrepancy may occur when the current aspect is *novel*: one which is unknown (i.e. not yet part of the model). In other words, it does not fit the child's model because it extends the model. A taxonomy of discrepancy is mapped in Figure 1, Right. Subsequent discussion of discrepancy in this paper excludes novelty because it does not involve any element of error by either learner or system.

Several examples of surprises and child reactions are described below, all of which are drawn from the data set and results described in Section 4.

- *Surprising event 1:* The child and Andy are completing a turn-taking activity. Unexpectedly, Andy walks off-screen and does not return. Andy is always programmed to stay onscreen during activities. After watching the side of the screen for a few seconds, the child makes a social reference to the researcher (i.e. gazing to seek information), and then looks back to the screen.

- *Surprising event 2:* Andy demonstrates a ball-sorting activity for the first time, putting multiple balls into the boxes of the same colour. After the child takes several turns, Andy tries to put a yellow ball in the red box (see Figure 2). It rolls off the top instead of dropping in. The child explicitly corrects Andy, pointing to the yellow box and excitedly shouting "Right here!"
- *Non-event 1:* In a flower-picking activity, Andy asks for the child's help and then indicates one of the three available flowers with a combination of gaze and pointing. The child tries touching all three flowers in turn: the flower indicated by Andy should fly into the basket when touched, but does not (i.e. apparently no available choice is correct). The child reacts by ceasing to touch the screen and leaning in to look very closely at Andy's face (i.e. social referencing; seeking information).
- *Non-event 2:* The Magic Garden fades in to start a new activity. There is an unusually long pause without Andy entering (i.e. long compared to previous activities), but background sound effects continue (i.e. the system is not frozen). The child reacts by asking the researcher "Where's Andy?"

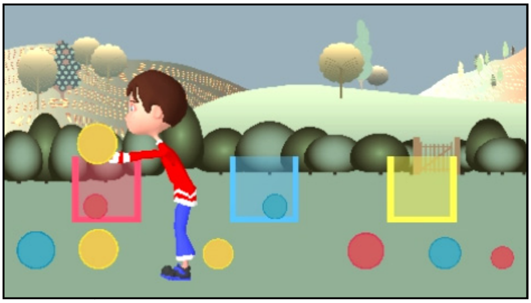


Fig. 2. Andy incorrectly sorts a ball (as described in Event example 2)

These qualitative examples illustrate the range of discrepancy-reaction pairs which were present in the current ECHOES dataset. The quantitative analysis described in Section 4 aimed to investigate their frequency and extent across the current participants, and whether the pattern varied across discrepancy types.

4 An Analysis of System-Side Errors in ECHOES

4.1 Method

Participants. The subset of ECHOES participants included in the discrepancy video analysis (n=8, 1 female, 7 male) were from two sites of the ECHOES summative evaluation study (in preparation). Each child had a previous diagnosis of an ASC by a paediatrician, child psychiatrist, or other professional. Their data was selected for additional analysis due to the children demonstrating at least phrase-language use and having sufficiently complete video samples (at least 30 minutes worth)[6]. All but one of the children in this group appear to have some degree of intellectual disability in addition to their ASC diagnosis, as evidenced by the discrepancy between their calendar ages (range= 5-8 yrs, mean= 6 yrs, 5 mo.) and verbal-mental ages[7] (VMA; range =2-5 yrs, 10mo., mean=3 yrs. 9 mo.).

Video Annotation. Each child's video samples were annotated for discrepancy-reaction pairs by the first author using ELAN [12], and in accordance with the categories described in Section 3 (see [10] for further details of the taxonomy and the annotation process). As noted in Section 3, discrepancy has a child-centred definition. As the child's understanding of the environment is generally private, with explicit statements of expectation or prediction relatively rare, observable child reactions are the only evidence for discrepancy detection. Thus, the unit of analysis is the *discrepancy-reaction pair*, not discrepancy alone.

The main source of information when inferring the presence of discrepancy-reaction pairs is knowledge of what the child has been exposed to in the environment (and how often). The annotator must consider the evidence a child might have about what is in the environment and how it "should" work. Surprises that could objectively be considered violations of the system's usual patterns (e.g. the VC making mistakes, or failing to appear) often signalled video sections that included discrepancy-reaction pairs, as did the introduction of a new activities or objects. Finding additional discrepancies involved observing the child's interaction with the environment, looking for cause-effect relationships between the system content and the child's behaviour.

Annotations noted whether child reactions were *initiations* (i.e. purposeful and spontaneous behaviours directed to a social partner), or *non-social reactions* (i.e. self-directed or undirected). The annotation recorded the target of the initiation (researcher or Andy) and also whether it was *primary* (the first reaction to that instance of discrepancy) or *secondary* (a subsequent initiation to the same instance of discrepancy). These categories aid in identifying reciprocal interactions about discrepancy.

Annotation data was exported from ELAN as tab-delimited text and further analysed in a standard spreadsheet program. Analysis focused on counting the instances in various categories, rather than seeking comparisons between participating children or between reactions to discrepancy as compared to other environmental events.

4.2 Results and Analysis

The spreadsheet analysis yielded 50 surprising event-reaction pairs and 71 non-event-reaction pairs from 347 minutes of video data. These totals include both primary

[6] The video data captured a variety of learning activities, as new material was introduced throughout. It consisted of three 15 minute samples from early, middle, and late sessions with the VE (45 minutes total per child). One participant had only 33 minutes of data due to missed sessions. Samples excluded non-analysable video (e.g. system crashes, child rest breaks) and learning activities in which the VC was not present.

[7] As calculated from their scores on the British Picture Vocabulary Scale (BPVS; [11]), a standardized measure of receptive language ability.

social reactions (initiations) and non-social reactions. Each child had between 9 and 22 pairs (mean=15.12, SD=4.12); it is encouraging that all children in the group both noticed and reacted to discrepancies, rather than reactions being concentrated in a few children only. Considering again the often severe social and communicative challenges that people with ASC may face (Section 2), perhaps the most notable result is that 54% of child reactions to surprising events and 69% of reactions to non-events were directed to the researcher or the VC (i.e. were initiations; mean= 61.05% of reactions). Furthermore, 33 out of these 121 discrepancy-reaction pairs (27.27%) formed the first in a *sequence* of child initiations about that same instance. Some of these sequences developed into reciprocal interactions, often verbal dialogues with the researcher.

Existing literature about the behavioural rigidity and insistence on sameness that frequently characterise ASC (see Section 2) suggests that participants might become severely emotionally dysregulated when they detect a discrepancy. However, there were few instances of obvious frustration and zero instances of the child "melting down" because the environment was breaking its own rules. The affect of the initiations was overwhelmingly positive or neutral with frequent smiles and laughter, as children appeared to find many of the system errors to be humorous.

5 Discussion

The results from the ECHOES video analysis are encouraging in and of themselves with respect to the specific user group, all of whom spontaneously engaged in social behaviours that are considered difficult for people with ASC [7], but are developmentally crucial. The current system, user group, and cause-and-effect type content are all undeniably specialised and may not be directly comparable to other teaching contexts, however, the underlying metacognitive process of discrepancy detection remains the same across contexts. It requires the learner to consider the current information or procedure in light of what he already knows (i.e. in comparison to his mental model), and to conclude that something "does not fit". Thus, the following sections use the current dataset as a starting point from which to theorise about discrepancy detection, system-side errors, and their potential as a pedagogic strategy.

5.1 Locating the Source of Errors

The discrepancy categories described and taxonomised in Section 3 (see also [10]) identify the *type* of mismatch between a mental model and the actual system/environment, but these categories are independent of the mismatch's *location*. In other words, they say nothing about whose error or misconception led to the mismatch. For example, several children using ECHOES requested help with unresponsive or "broken" digital objects that were in fact functional, but unable to detect their inappropriate touch screen actions (e.g. scratching or hitting). This was not a problem with the system, but with the child's mental model of the object (or rather, the actions by which it could be affected). From the child's view, there was a discrepancy between the action's expected result and the object's failure to respond (an example of a non-event).

For any given piece of system content for which the learner has a mental model, there are four possible combinations of errors and correct knowledge, only some of which afford discrepancy-detection. Table 1 explains these combinations. The location of an error matters when determining a pedagogic strategy. The end goal is usually to reach state A, alignment of learner and system knowledge. Most pedagogic strategies work towards state A from state C, learner-side errors, with the expert applying correct knowledge in order to support the learner in correcting the item. However, as the current video data illustrates, system-side errors (state B) can also galvanise learners to metacognitively reflect on their models, locate errors, and even offer correction (i.e. move toward state A). Correcting errors in teachable agent system (e.g. [13]) appears to have elements of both B and C, because the learner corrects "the agent's" mistakes, which are apparently external to the user (i.e. a system-side or at least system-*like* error, as in B), but she is actually reflecting on and amending her *own* externalized domain knowledge (a learner-side error, as in C). *Compound Errors* (state D) will not necessarily lead to this constructive metacognition and resolution, as the learner and the system may not be in a position to correct one another. Table 1 supports the taxonomy of discrepancy types briefly outlined in Section 3 and further expanded in [10] as, taken together, they provide a high-level description of a discrepancy's type and location. Deliberate system-side errors or erroneous examples appear to still be an "emerging" area for educational technologies, and while unlikely to be applicable to all domains, may prove to be a useful lens through which to describe and compare work in this area.

Table 1. Possible locations of discrepancy between learners' mental models of some kind *X* and a specific instance *x* in the system

	System behaves **correctly or consistently**[8] on *x*.	System behaves **incorrectly or inconsistently** on *x*.
Learner's mental model **correct** regarding *X*.	**A.** Learner-domain alignment (no error; no discrepancy to detect)	**B.** System-side error (Learner may detect error as a source of discrepancy)
Learner's mental model **incorrect** regarding *X*.	**C.** Learner-side error (Learner may detect error via metacognition or may require system's direction)	**D.** Compound error (2 sources of discrepancy, 4 possible outcomes with respect to detection/ non-detection)[9]

[8] Behaviour is in accord with domain "facts" or "rules", (however represented), or behaviour consistent with the system's own procedures (outside of the targeted teaching material).

[9] One discrepancy may result from the learner's error and another from the system's. The outcomes depend on whether or not the learner detects either of those errors.

5.2 Extending System-Side Errors: Domains, Users, Unanswered Questions

As other authors have already acknowledged [2, 3], a long list of foundational questions remain to be resolved before any system could be designed that employs deliberate errors and/or facilitates discrepancy detection in a truly adaptive way. Although the current research provides a framework in which to better understand the nature of such errors, further research is needed to address general questions regarding when and how often to deliberately introduce system-side errors, and whether or not they are equally appropriate for all types of learners or all levels of domain proficiency. Instead of providing answers, the current work is an example of how the general strategy of system-side errors motivating metacognition could be successful in a very different type of situation than has previously been investigated, or to which adaptive systems are most often applied. The main areas of difference are as follows:

- Learners' young age and significant additional support needs
- Exploratory system, not focused on explicit problem-solving or content-rehearsal
- Errors are "unannounced" rather than presented as a specific exercise, example, or teaching opportunity. This is of course due to the fact that the system-side errors in ECHOES were not a deliberate design decision; see Section 1.
- Domain content is non-propositional (social communication skills).
- Errors are also non-propositional, and constitute disrupted cause-and-effect relationships, or alterations of sensory or temporal aspects of the environment

The ECHOES video analysis illustrates a very different case of system-side errors than those in existing mathematics-focused work (e.g. [2, 3]), but arguably draws on the same underlying metacognitive processes of model-comparison and discrepancy detection. Presenting learners with deliberate errors may be a more widely applicable strategy than it initially appears.

6 Conclusion

In summary, the ECHOES dataset illustrates that occasional system-side errors can motivate children to spontaneously reflect on their mental model of an environment, and to spontaneously articulate information to social partners about discrepancies between their models and the system, and how these might be remedied. This appears to be an equivalent metacognitive process to learners correcting their own errors. In the context of ECHOES, these system-side errors brought clear benefits for learners, suggesting that use of errors to promote metacognition and content practice can be usefully extended to very different content and user groups than have previously been investigated. The taxonomy of discrepancy types and table of error locations presented in this paper attempt to abstract away from the ECHOES context, and suggest a means of describing discrepancies that may be useful in comparing and synthesizing work in this emerging area of educational technologies and pedagogic strategy.

Acknowledgements. This research was undertaken as part of a PhD thesis by the first author, funded by the University of Edinburgh and the Scottish Informatics and Computer Science Alliance. It includes materials from the ECHOES project (funded by ESRC/EPSRC TEL; RES-139-25-0395). The ECHOES environment was designed and developed by the ECHOES team, including the authors (see www.echoes2.org), with contributions from various stakeholders. Special thanks to the participating schools, staff, and children. Their time and enthusiasm made this work possible.

References

1. Du Boulay, B., Luckin, R.: Modelling human teaching tactics and strategies for tutoring systems. Int. J. of Artificial Intelligence in Education 12(3), 235–256 (2001)
2. Tsovaltzi, D., McLaren, B.M., Melis, E., Meyer, A.K.: Erroneous examples: effects on learning fractions in a web-based setting. Int. J. Technology Enhanced Learning 4(3/4), 191–230 (2012)
3. McLaren, B.M., et al.: To Err Is Human, to Explain and Correct Is Divine: A Study of Interactive Erroneous Examples with Middle School Math Students. In: Ravenscroft, A., Lindstaedt, S., Kloos, C.D., Hernández-Leo, D. (eds.) EC-TEL 2012. LNCS, vol. 7563, pp. 222–235. Springer, Heidelberg (2012)
4. Porayska-Pomsta, K., Frauenberger, C., Pain, H., Rajendran, G., Smith, T., Menzies, R., Foster, M.E., Alcorn, A., Wass, S., Bernadini, S., Avramides, K., Keay-Bright, W., Chen, J., Waller, A., Guldberg, K., Good, J., Lemon, O.: Developing technology for autism: an interdisciplinary approach. Personal and Ubiquitous Computing 16(2), 117–127 (2011)
5. Alcorn, A., et al.: Social communication between virtual characters and children with autism. In: Biswas, G., Bull, S., Kay, J., Mitrovic, A. (eds.) AIED 2011. LNCS, vol. 6738, pp. 7–14. Springer, Heidelberg (2011)
6. Flavell, J.H.: Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. American Psychologist 34(10), 906–911 (1979)
7. DSM-IV: Diagnostic and statistical manual of mental disorders. American Psychiatric Association Washington, DC (1994)
8. Rajendran, G.: Virtual environments and autism: a developmental psychopathological approach. J. of Computer Assisted Learning (2013), doi:10.1111/jcal.12006
9. Foster, M., Avramides, K., Bernardini, S., Chen, J., Frauenberger, C., Lemon, O., Porayska-Pomsta, K.: Supporting Children's Social Communication Skills through Interactive Narratives with Virtual Characters. In: Proceedings of the International Conference on Multimedia, pp. 1111–1114 (2010)
10. Alcorn, A.M., Pain, H., Good, J.: Discrepancies in a Virtual Learning Environment: Something "Worth Communicating About" for Young Children with ASC? In: International Conference on Interaction Design and Children (IDC 2013), New York (in Press, 2013)
11. Dunn, L.M., Dunn, D.M., Whetton, C.W., Burley, J.: The British Picture Vocabulary Scale, 2nd edn. NFER Nelson, Windsor (1997)
12. Max Planck Institute for Psycholinguistics. ELAN Linguistics Annotator, version 4.4.0. Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands (2012), http://tla.mpi.nl/tools/tla-tools/elan/
13. Biswas, G., Leelawong, K., Schwartz, D., Vye, N.: Learning by teaching: A new agent paradigm for educational software. Applied Artificial Intelligence: An International Journal 19, 363–392 (2005)